The complete set of tRNA species in *Nanoarchaeum equitans*

Lennart Randau[a], Michael Pearson[b], Dieter Söll[a,c,*]

[a] *Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA*
[b] *Department of Molecular, Cellular and Developmental Biology, Sinsheimer Laboratories, UCSC, Santa Cruz, CA 95064, USA*
[c] *Department of Chemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA*

Received 22 April 2005; accepted 25 April 2005

Available online 3 May 2005

Edited by Lev Kisselev

Abstract The archaeal parasite *Nanoarchaeum equitans* was found to generate five tRNA species via a unique process requiring the assembly of seperate 5′ and 3′ tRNA halves [Randau, L., Münch, R., Hohn, M.J., Jahn, D. and Söll, D. (2005) *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5′- and 3′-halves. Nature 433, 537–541]. Biochemical evidence was missing for one of the computationally-predicted, joined tRNAs designated as $tRNA^{Trp}$. Our RT-PCR and sequencing results identify this tRNA as $tRNA^{Lys}$ (CUU) joined at the alternative position between bases 30 and 31. We show that the intron-containing $tRNA^{Trp}$ was misidentified in the initial *Nanoarchaeum equitans* genome annotation [E. Waters et al. (2003) The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. Proc. Natl. Acad. Sci. USA 100, 12984–12988]. Along with a previously unidentified joined $tRNA^{Gln}$ (UUG), *Nanoarchaeum equitans* exhibits 44 tRNAs and is enabled to read all 61 sense codons. Features unique to this set of tRNA molecules are discussed.

Keywords: tRNA splicing; tRNA maturation; Intron

1. Introduction

The sequencing of the small genome of the hyperthermophilic parasite *Nanoarchaeum equitans* raised some questions since the genes for four essential transfer RNA species appeared to be missing [2]. Recently, this mystery was solved by the finding that *N. equitans* is the only known organism to assemble full-length tRNAs from separate genes encoding their 5′- and 3′-halves [1]. The assembly of four tRNA species ($tRNA^{Glu}$ (CUG), $tRNA^{Glu}$ (UUG), $tRNA^{His}$ and $tRNA_i^{Met}$) from nine tRNA halves was shown biochemically, but no sequence information was obtained for four putative tRNA halves. It was predicted that the mature *N. equitans* tRNAs are joined one base after the anticodon-adjacent position 37, the location of most tRNA introns which led to identification of a "missing" $tRNA^{Trp}$. We reinvestigated possible tRNAs joined at alternative solutions. Our new sequencing results led to the reannotation of three tRNA genes and provide *N. equitans* with a complete set of tRNAs.

2. Materials and methods

2.1. Cell cultivation and tRNA isolation

N. equitans cells were grown in a 300 l fermenter in a simultaneous culture with *Ignicoccus* sp. and purified by gradient centrifugation as described [3]. Total tRNA was prepared by SDS-lysis of the cell pellet, phenol/chloroform extraction and MonoQ anion-exchange chromatography as described [1].

2.2. Reverse transcription and sequencing

Total tRNA from *N. equitans* was reverse transcribed with Thermoscript reverse transcriptase and the resulting cDNA was PCR amplified with Platinum *Taq* DNA polymerase (Invitrogen) according to the manufacturer's directions. The tRNA template and primers were denatured at 100 °C for 5 min and cooled on ice for 5 min to facilitate specific annealing. PCR products were cloned with the pCR-2.1-TOPO cloning kit (Invitrogen). Plasmids were sequenced at the W.M. Keck Facility. The following oligonucleotides were used for PCR amplification of the indicated full-length tRNAs: ($tRNA^{Lys}$) 5′-GGGCCGGTAGCTCAGCCTGG-3′ and 5′-CGGGCCGGCGGGGATTCGAACC-3′, ($tRNA^{Gln}$) 5′-AGCCCCGTGGTGTAGCGGC-3′ and 5′-TAGCCCCGCCCGGATTCGAACC-3′, ($tRNA^{Trp}$) 5′-GGGGCCGTAGCTCAGCCAGGCAG-3′ and 5′-TGGGGCCGGGGGGATTCGAACC-3′ ($tRNA^{Met}$) 5′-GCCGCCGTAGCTCAGCGG-3′ and 5′-TGCCGCCGGCGGGATTCGAACC-3′.

3. Results and discussion

We reverse transcribed total tRNA from *N.equitans* with primers complementary to the 3′ ends of the interrupted tRNAs. The resulting cDNAs were amplified by PCR with an isotype-specific set of oligos described in Section 2 (Fig. 1). The sequencing results for the joined tRNA previously designated as $tRNA^{Trp}$ stimulated our interest. This tRNA was found to be $tRNA^{Lys}$ (CUU) with the two tRNA halves joined at an alternative position between nucleotides 30 and 31 (Fig. 2). This finding challenges the view that tRNA halves are specifically assembled at position 37 and led to the reinvestigation of two additional tRNA halves identified by the long reverse complementary intervening sequence characteristic for all split tRNAs. We determined that these tRNA halves are joined at position 32 and a previously unidentified $tRNA^{Gln}$ (UUG) is formed (Fig. 2). This isoacceptor features the same recognition elements for glutaminyl-tRNA synthetase present in the uninterrupted isoacceptor $tRNA^{Gln}$ (CUG) which includes the unique A1-U72 basepair [4]. The two split $tRNA^{Glu}$ isoacceptors

*Corresponding author. Fax: +1 203 432 6202.
E-mail address: soll@trna.chem.yale.edu (D. Söll).

Abbreviations: BHB, bulge helix bulge; RT-PCR, reverse transcriptase PCR

0014-5793/$30.00
doi:10.1016/j.febslet.2005.04.051

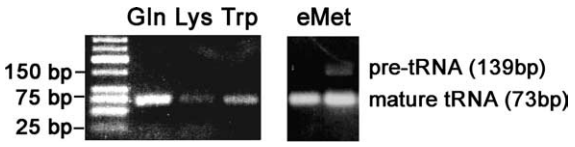

Fig. 1. RT-PCR results. The RT-PCR products obtained for $tRNA^{Gln}$ (UUG) (Gln), $tRNA^{Lys}$ (CUU) (Lys) and $tRNA^{Trp}$ (Trp) were separated on a 3% ethidium bromide-stained agarose gel. The fragments of a low molecular weight DNA ladder are indicated for size comparison. RT-PCR of elongator $tRNA^{Met}$ (eMet) yielded two amplified fragments. The individual bands were gel-excised, cloned and sequenced. The upper band was shown to be the intron-containing pre-$tRNA^{Met}$ and the lower band the spliced mature $tRNA^{Met}$ annotated in [2].

and both $tRNA^{Gln}$ isoacceptors exhibit a deletion of base 47 (Fig. 3A). The bona fide $tRNA^{Trp}$ from *N.equitans* consequently is not assembled from separate tRNA halves but rather is the product of an intron-containing tRNA gene which was previously misidentified as $tRNA^{Ser}$ during the initial genome annotation using tRNAScan-SE [5]. The position of the intron was proposed to be between bases 37 and 38, thus predicting a second *cis*-spliced $tRNA^{Ser}$ (CGA) isoacceptor. However, this $tRNA^{Ser}$ does not contain the signature elements required for recognition by seryl-tRNA synthetase, namely the extended variable loop [6]. We found that the 13 nucleotide long intron is removed between the base 30 and 31 and that the resulting mature $tRNA^{Trp}$ features the anticodon CCA and the discriminator base A73 recognized by tryptophanyl-tRNA synthetase [7] (Fig. 2). The new annotations describe six split tRNAs and four *cis*-spliced tRNAs, finally providing *N. equitans* with a set of 44 non-reduntant tRNAs able to read all 61 sense codons (Fig. 3A).

Two features unique to the tRNA set of *N. equitans* should be noted. Firstly, *N. equitans* is the only known archaeon displaying a $tRNA^{Ile}$ (UAU) anticodon which might correlate with the absence of a protein coding gene required to modify the CAU anticodon. Secondly, the initiator tRNA exhibits a C1–G72 base pair whereas a A1–U72 base pair is the consensus element in all other archaeal and eukaryotic initiator tRNAs. Our analysis of boxA promoter motifs [1] usually found in the distance of 26 nucleotides towards the tRNA gene suggests the possibility that the acceptor stem might be extended by one base pair which would allow the presence of a consensus A–U pair at its end. Thus, we propose that the *N. equitans* $tRNA_i^{Met}$ has an 8 bp acceptor stem.

We initially observed tRNA splicing of *N. equitans* elongator $tRNA^{Met}$ when reverse transcriptase PCR (RT-PCR) experiments amplified the mature, processed tRNA as well as the primary transcript containing the 66 bp long intron. In order to investigate the mechanism of tRNA splicing in *N. equitans* we took a closer look at the sequences around the tRNA split sites. Interestingly, we noted the possible formation of relaxed Bulge-Helix-Bulge (BHB) motifs for the *cis*-spliced tRNAs (Fig. 3A) as well as for the joined tRNA halves (Fig. 3C). This correlation suggests the possibility of a *trans*-splicing mechanism in the assembly of the tRNA halves. In this case a splicing endonuclease would be required to recognize the proposed relaxed BHB-motifs. Possible evidence for biological significance of these relaxed motifs is the observation that a large number of tRNA introns in Crenarchaeota can only be folded into non-standard BHB motifs, termed hBH or HBh′ [8]. Here, only the central 4-bp helix H, one bulge B and a flanking helix on either side (h or h′) can be identified. Experimental investigations will be needed to define the biochemical mechanism of tRNA *trans*-splicing. The overall similarity of the proposed splicing motifs point out an appealing connection between the mechanism of assembly of tRNA halves and the presence of introns in tRNAs.

A function for divided tRNA genes is not immediately obvious. However, the observation that site-specific integration by archaeal viruses and conjugative plasmids may occur exclusively at tRNA genes [9] leads us to consider that an adaptive value may lie in providing resistance to the integration of mobile DNA elements. As *N. equitans* has virtually no metabolic or biosynthetic genes [2], the added burden of simply replicating the DNA of an integrated element would be expected to slow growth considerably. While a typical tRNA intron might also function to prevent integration, extreme genome reduction would favor its deletion when the selection pressure of integration were not present.

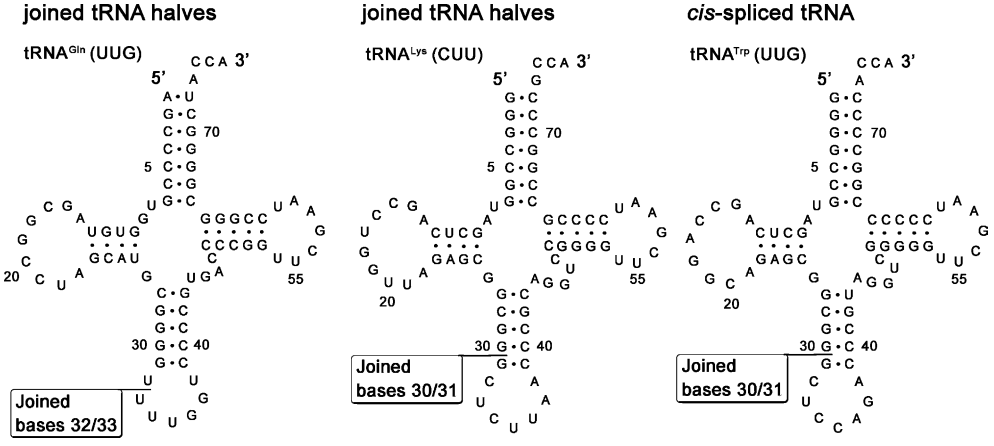

Fig. 2. Sequencing results for processed tRNAs of *N. equitans*. Secondary structures for the processed tRNAs obtained by RT-PCR and sequencing are displayed. Numbering of the positions is according to Sprinzl et al. [10].

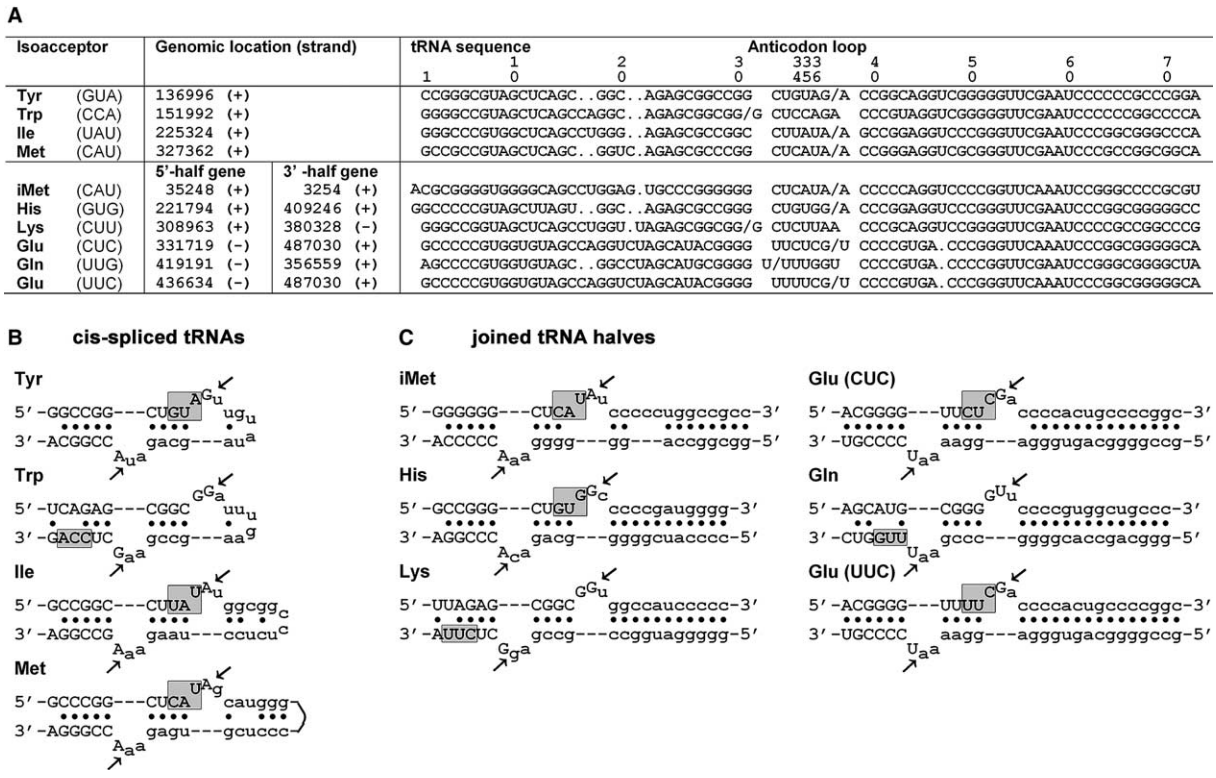

Fig. 3. The set of processed tRNAs of *N. equitans*. (A) Alignment of tRNAs with the genomic position indicated. The tRNA halves are joined at the position indicated by a slash. Possible secondary structures of relaxed Bulge-Helix-Bulge motifs for (B) *cis*-spliced tRNAs and (C) joined (*trans*-spliced) tRNA half genes are given. The individual anticodons are boxed and the splice sites are indicated by arrows. The tRNA sequence is written in uppercase letters and the excised intervening sequences are written in lowercase letters.

In organisms where a chromosome rearrangement divided the intron, the resistance to integration would be stabilized. Upon repeated exposure to an integrative element, *trans*-spliced tRNA genes could become fixed in the population.

Acknowledgements: We thank J. Yuan and J. Sabina for critically reading the manuscript. This work was supported by grants from the National Institute of General Medical Sciences and the Department of Energy.

References

[1] Randau, L., Münch, R., Hohn, M.J., Jahn, D. and Söll, D. (2005) *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5′- and 3′-halves. Nature 433, 537–541.
[2] Waters, E. et al. (2003) The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. Proc. Natl. Acad. Sci. USA 100, 12984–12988.
[3] Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C. and Stetter, K.O. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature 417, 63–67.
[4] Freist, W., Gauss, D.H., Ibba, M. and Söll, D. (1997) Glutaminyl-tRNA synthetase. Biol. Chem. 378, 1103–1117.
[5] Lowe, T.M. and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
[6] Korencic, D., Polycarpo, C., Weygand-Durasevic, I. and Söll, D. (2004) Differential modes of transfer RNA^{Ser} recognition in *Methanosarcina barkeri*. J. Biol. Chem. 279, 48780–48786.
[7] Guo, Q. et al. (2002) Recognition by tryptophanyl-tRNA synthetases of discriminator base on $tRNA^{Trp}$ from three biological domains. J. Biol. Chem. 277, 14343–14349.
[8] Marck, C. and Grosjean, H. (2003) Identification of BHB splicing motifs in intron-containing tRNAs from 18 archaea: evolutionary implications. RNA 9, 1516–1531.
[9] She, Q., Shen, B. and Chen, L. (2004) Archaeal integrases and mechanisms of gene capture. Biochem. Soc. Trans. 32, 222–226.
[10] Sprinzl, M. and Vassilenko, K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. 33, D139–D140.